Pergamon

PII: S0042-6989(96)00207-6

Vision Res., Vol. 37, No. 6, pp. 745–755, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Motion-reversal Reveals Two Motion Mechanisms Functioning in Scotopic Vision

TATSUTO TAKEUCHI,*†‡ KAREN K. DE VALOIS†

Received 5 February 1996; in revised form 20 May 1996

We studied scotopic motion mechanisms, using a two-frame sinusoidal grating separated by various ISIs equated for mean luminance level. Perceived direction of displacement varied with both ISI and luminance. As luminance decreased, apparent motion reversal disappeared. This is predicted by a first-order motion model if the underlying temporal impulse response function varies from biphasic under photopic conditions to monophasic under scotopic conditions. Performance at long (but not short) ISIs depends upon stimulus contrast, suggesting there is also a scotopic feature-tracking mechanism. With isoluminant and high spatial frequency gratings, where the temporal impulse response function is monophasic, no motion reversal was observed. © 1997 Elsevier Science Ltd. All rights reserved.

Motion Scotopic vision Motion energy Feature tracking

INTRODUCTION

Under some conditions, the detection and discrimination of motion can be well understood by considering the operation of presumed low-level detectors that act essentially as spatiotemporal orientation detectors (Adelson & Bergen, 1985; van Santen & Sperling, 1984, 1985; Watson & Ahumada, 1985; Burr *et al.*, 1986). These mechanisms are called first-order motion detectors (e.g., Cavanagh & Mather, 1989). In other circumstances, it may be necessary to consider a higher-order mechanism that examines the change in position of identifiable pattern features over time (Ullman, 1979; Dawson, 1991). Here, the underlying mechanism is assumed to solve the correspondence problem, in which a given element in one image (Frame 1) is identified as and matched to the same element in the next image (Frame 2). Because it is intrinsically difficult to solve the correspondence problem, a major advantage associated with first-order motion detectors is that this task is avoided in their formulation (Adelson & Bergen, 1985). The motion of so-called second-order stimuli, patterns that are defined by spatiotemporal variations other than those of luminance or color, can be accomplished by the extraction of spatiotemporal orientation by a mechanism with an early non-linear transformation such as rectification (Chubb & Sperling, 1988). Psychophysical research suggests that these three kinds of motion detection mechanisms are implemented in the human visual system (Smith, 1994; Lu & Sperling, 1995).

Although an extraordinary amount of research during the last few years has addressed visual motion analysis, virtually all has concerned motion perception under photopic conditions. Vision under low light conditions is assuredly important and in some respects well studied, but we know remarkably little about the perception of motion at scotopic adaptation levels (Snowden *et al.*, 1995). Here we raise the following questions: Does scotopic motion perception depend upon a first-order luminance motion system that extracts spatiotemporal orientation? Does it rely upon the sequential comparison of positions over time? Are both of these kinds of motion mechanisms involved?

When considering the analysis of visual motion, the temporal sensitivity of the system is of particular interest. The shape of the luminance temporal impulse response function§ under photopic conditions, as estimated from psychophysical studies, is biphasic, with one positive lobe followed by a second negative lobe (see Ikeda, 1986 for a review). However, as the adapting luminance level decreases, temporal sensitivity changes systematically (Kelly, 1971; Swanson *et al.*, 1987). Swanson *et al.*

*Present address: Information Science Research Laboratory, NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa 243-01, Japan.
†Psychology Department and Group in Vision Science, 3210 Tolman Hall #1650, University of California at Berkeley, Berkeley, CA 94720-1650, U.S.A.
‡To whom all correspondence should be addressed.

§Under the assumption of linearity, the modulation transfer function as a function of temporal frequency is the Fourier transform of the temporal impulse response function (e.g., Watson, 1986). The impulse response function allows one to predict how the system will respond to an arbitrary input by convolving the input and the impulse response function.

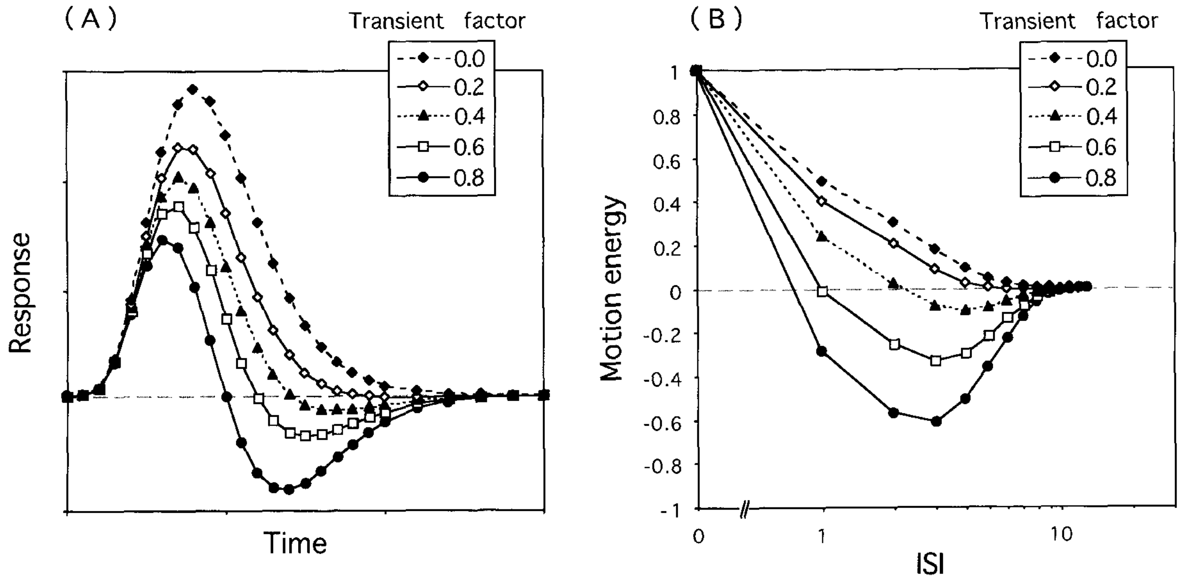

FIGURE 1. (A) Model temporal impulse response functions computed from Equation (1). The transient factor B varies from 0.8 (filled circles) to 0.0 (filled diamonds) in five steps. Other parameters of Equation (1) are $n = 6$ and $k = 1.5$. (B) The output of the motion energy model as a function of ISI of two-frame sine-wave gratings. Positive values of Y represent motion energy corresponding to displacement of the second frame in the direction of $\pi/2$, and negative values represent reversed motion, corresponding to displacement in the direction of $3\pi/2$. The value at ISI = 0.0 is normalized to 1.0. The transient factor B in Equation (1) was varied from 0.8 (filled circles) to 0.0 (filled diamonds).

(1987) derived estimates of the temporal impulse response function for the human visual system under photopic and scotopic conditions, examining both sensitivity to temporal flicker and the relationship between pulse duration and modulation sensitivity. The photopic and scotopic temporal impulse response functions differ in certain respects. First, the latency from response onset to the peak of the first (positive) deviation increases as luminance level decreases, and the duration of the first deviation increases. Second, for lower light levels, the magnitude of the second (negative) deviation falls and effectively disappears at the lowest (scotopic) levels. The disappearance of the negative lobe will be of particular interest in the discussion that follows.

Several models proposed for luminance-based (or first-order) motion detection contain temporal filters at the front end (Adelson & Bergen, 1985; Marr & Ullman, 1981; van Santen & Sperling, 1984, 1985; Watson & Ahumada, 1985). The oriented spatiotemporal filters proposed by Adelson and Bergen (1985), for example, are constructed from two quadrature pairs made from four spatiotemporal separable filters, which are themselves the products of the spatial and temporal impulse response functions. Their model includes the biphasic temporal impulse response function characteristic of the photopic system, not the monophasic scotopic function.

We have investigated the hypothesis that a first-order motion detector based upon a monophasic temporal impulse response function operates under scotopic conditions. To test this hypothesis, we took advantage of an interesting visual motion illusion (Braddick, 1980). When a single pattern is presented in first one position, then another, an observer may experience a strong sensation of motion, traditionally called apparent motion, if the temporal and spatial separations lie within appropriate ranges. If the two presentations are separated by a brief inter-stimulus interval (ISI) within a certain range, however, and if the interval is filled with a blank screen equated in space-averaged luminance to the pattern display, then the apparent direction of motion will be reversed (Braddick, 1980; Boulton & Baker, 1993). Shioiri and Cavanagh (1990) and Pantle and Turano (1992) have explained this very counterintuitive phenomenon by assuming that an underlying mechanism with a biphasic temporal impulse response function feeds into the responsible motion detector. The finding by Pantle and Turano (1992) that the motion reversal does not occur when a contrast-modulated pattern is used as an input stimulus supports the suggestion that a first-order motion detector contributes to this phenomenon. This raises an interesting possibility. If the temporal impulse response function for the scotopic system is monophasic (or very nearly so), first-order motion detectors like those modeled by Adelson and Bergen (1985) would have space–time receptive fields that differ from those of the photopic system. A corollary is that the reversed motion illusion described above should fail to appear when the light level is very low. Of course, this prediction assumes that the scotopic system contains spatiotemporal filtering mechanisms that combine to form first-order motion sensors, an assumption for which we have little evidence. One of our aims, thus, is to determine whether scotopic motion perception can be modeled by first-order motion mechanisms, or whether it must rely upon some other kind of motion sensor, such as a feature-tracking system (Ullman, 1979).

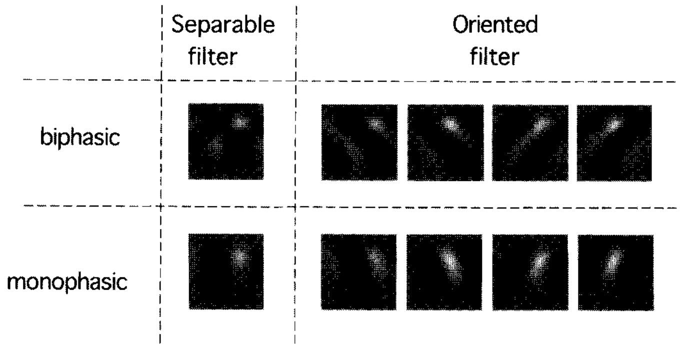

FIGURE 2. An x-t description of spatiotemporal separable filters and oriented filters in which the underlying temporal impulse response function is biphasic (upper panel) or monophasic (lower panel). To construct the spatiotemporal separable filters, the temporal impulse response function was taken from Equation (1), where $n = 6$ and $k = 1.5$. In the upper panel, the temporal impulse response function used was biphasic, with $B = 1.0$. In the lower panel, $B = 0.0$, producing a monophasic temporal impulse response function. The spatial impulse response function was a cosine Gabor function. Spatiotemporal oriented filters were constructed from the separable filters shown and other similar ones in which the spatial Gabor was in sine phase and/or $n = 9.0$ in Equation (1) (i.e., the quadrature pair of temporal impulse response functions), following the formulation of Adelson and Bergen (1985).

PREDICTION FROM A FIRST-ORDER MOTION MODEL

We implemented the motion energy model of Adelson and Bergen (1985), in an attempt to predict results from our psychophysical experiments. Equation (1) describes the temporal impulse response function used by Adelson and Bergen (1985) and by Emerson *et al.* (1992), which was based on temporal flicker sensitivity data from Robson (1966).

$$f(t) = (kt)^n \exp(-kt)[1/n! - B(kt)^2/(n+2)!] \quad (1)$$

Here, n is the tuning width in the frequency domain, and k is the filter center frequency (Nowlan & Sejnowski, 1994). B defines the weighting of the negative phase relative to the first positive phase, which can be called a transient factor (Watson, 1986). Figure 1(A) illustrates this model of the temporal impulse response function.

As the transient factor B is decreased, the negative phase disappears and the positive peak shifts to the right. These characteristics are qualitatively similar to the changes in temporal impulse response functions derived by Kelly (1971) and by Swanson *et al.* (1987) as luminance falls. Decreasing the transient factor corresponds to a reduction of low frequency attenuation in the temporal frequency domain. Thus, as the transient factor falls, the temporal MTF becomes increasingly low pass. We derived a quadrature pair of filters from the same temporal impulse response function. The spatial filters were two-dimensional sine and cosine phase Gabor functions, constituting a spatial quadrature pair. Gabor functions approximate the receptive fields of simple cells in V1 (De Valois & De Valois, 1988).

The stimulus used for both simulation and experiment was a two-frame sine-wave grating with various ISIs. The same stimulus was used by Pantle and Turano (1992) and Strout *et al.* (1994). Grating phase was chosen to be random for the first frame. In the second frame, the grating was shifted either up or down by a displacement of $\pi/2$. With ISI $= 0$, this displacement elicits a strong motion signal at near-threshold (Nakayama & Silverman, 1985) and suprathreshold (e.g., Ramachandran & Cavanagh, 1987; Turano & Pantle, 1985) contrast levels. As Pantle and Turano (1992) showed, the same stimulus also elicits a strong perception of reversed motion, that is, motion in the direction corresponding to a $3\pi/2$ displacement, at some ISI values.

To compute the output of a first-order motion sensor, a two-dimensional image with 256 gray levels representing an x–t stimulus plot was prepared. The four spatiotemporal separable filters used in the next stage are the products of two spatial and two temporal impulse response functions, as described by Adelson and Bergen (1985). These four spatiotemporal separable filters were convolved with the image. Further computation allows one to extract a motion energy signal (for details see Fig. 18(b) of Adelson & Bergen, 1985). The motion energy signal contains directional information corresponding to orientation in the x–t plane in a phase-independent manner.

To illustrate the effects of shifting from a biphasic temporal filter to a monophasic temporal filter, we show in Fig. 2 two spatiotemporal separable filters (cosine spatial phase) based on two different temporal impulse response functions (biphasic and monophasic). Note that these do not reflect separable filters constructed from a temporal quadrature pair. The center spatial frequency was the same as that of the stimulus. Figure 2 also shows two sets of four spatiotemporal oriented filters (Adelson & Bergen, 1985), for each (biphasic and monophasic) temporal impulse response function. Because we used a one-dimensional sine-wave grating as a stimulus, we projected the three-dimensional plane (x, y for the spatial dimension, t for the temporal dimension) to a two-dimensional plane (x for the spatial dimension, t for the temporal dimension) by averaging along the y-axis, which is parallel to the spatial filters' preferred orientation. The resulting field size of the spatiotemporal separable filters and oriented filters was arbitrarily set to 20 (spatial) $\times$ 20 (temporal) pixels, and the stimulus size in the x–t domain, to 128×128 pixels. The frame duration of each of the two sine-wave presentations was set to 40 pixels in the time domain. The output of the motion energy detector was displayed in a 128×128 image plane. Motion energy was defined as the space–time average of the entire output display.

Figure 1(B) shows the motion energy calculated for various values of the transient factor B. Values of Y greater than zero represent motion energy corresponding to displacement in the direction of $\pi/2$, and values below zero represent reversed motion, corresponding to displacement in the direction of $3\pi/2$. Each curve was normalized to 1.0 at ISI = 0. The results confirm that the presence of a negative lobe in the temporal impulse response function shown in Fig. 1(A) is critical for producing motion energy in the reversed (i.e., $3\pi/2$) direction, as discussed by Shioiri and Cavanagh (1990) and Strout *et al.* (1994). This leads to the prediction that the apparent motion reversal will disappear if the two-frame display is observed under scotopic conditions. Note also [in Fig. 1(B)] that the negative peak shifts to the right and becomes smaller as the transient factor decreases. The rightward shift suggests that as adapting luminance decreases, the ISI at which the reversed motion illusion is strongest should become increasingly longer. The decreasing amplitude of the negative lobe leads to the prediction that the absolute probability of seeing reversed motion should decrease systematically as luminance falls. For very long ISIs, motion energy reaches zero amplitude, indicating that no motion energy signal is generated in this time range. Strout *et al.* (1994) showed that when the ISI is short, the prediction from a motion energy computation (like that of Adelson & Bergen, 1985) is consistent with the output of their Phase II model. They did not report results for longer ISIs, however, so it is unclear whether direction discrimination reaches chance level for the two-frame sine-wave stimulus at longer ISIs.

EXPERIMENT 1

The purpose of Experiment 1 was to test the predictions described above, based on the assumption that a first-order motion mechanism operates at scotopic levels.

Subjects

Two subjects (LF, TT) participated in Experiments 1 and 2. LF was unaware of the purpose and ongoing results of the experiment. TT is one of the authors. A second naïve subject (AL) participated in Experiment 3. All had normal or corrected-to-normal vision and were between 22 and 30 years old.

Apparatus

Stimuli were generated on a Sun3/160 workstation with a TAAC graphics accelerator and displayed on a 16 in. RGB monitor (Sony GDM1604). The frame rate of the monitor was 66 Hz, with spatial resolution of 1152×900 pixels and gray-level resolution of 8 bits. The monitor was calibrated with a Minolta photometer, and its output was linearized (gamma corrected) under software control. Especially for stimuli displayed under photopic conditions, spatial dithering was used to produce very low contrasts. This was invisible to the observers. For all experiments using luminance-varying stimuli, the space-averaged chromaticity (CIE 1931) of the display was $x = 0.305$, $y = 0.323$. Subjects observed the display through a 2 mm artificial pupil, with head position maintained by a bite bar mounted on an XYZ translator. Viewing distance was 115 cm. The mean adapting level was varied by placing neutral density filters just distal to the artificial pupil. The average luminance level of the display was 25.0 cd/m^2, or 78.5 photopic td (1.9 log photopic td). The room was darkened and light shielded, with no other source of illumination present. We also used adapting levels of 7.85 td (0.9 log photopic td), 0.785 td (-0.1 log photopic td), and 0.0785 td (-1.1 log photopic td), respectively. We assume that only the scotopic system is active under the lowest adapting level (Hecht & Shlaer, 1936; Hood & Finkelstein, 1986; McCourt, 1990). Subjects initially dark adapted for 25 min prior to the task, and the experiment always started at the lowest adapting level.

Contrast sensitivity measurements

To equate in terms of multiples of threshold contrast for the different adapting levels, we measured contrast sensitivity for direction discrimination of the two-frame sine-wave gratings. A horizontal sine-wave grating was displayed in a 4.0 (H) $\times$ 6.0 (V) deg rectangular window centered in the display. Only the stimulus window was illuminated; the remainder of the screen was dark ($<$0.01 cd/m^2). The horizontal edges of the stimulus were tapered by a Gaussian function with sigma = 1.0 deg. The vertical edges were not tapered. The grating presented in the first frame was phase-shifted either upward or downward by $\pi/2$. No ISI was interposed; the pattern changed abruptly between two presentations of the 66 Hz display. The duration of each frame was 500 msec with a rectangular temporal window. We used a two-alternative, temporal forced-choice procedure. In one of two intervals, the motion was upward; in the other interval, it was downward. We refer to the true direction of motion as being the direction of the shortest path, that is, the direction in which the displacement was equivalent to $\pi/2$. The subject, by pressing one of two buttons, indicated which interval contained the upward motion. The two intervals were separated by a 1 sec blank field of the same space-averaged luminance, and the onset of each interval was marked by an auditory cue. No feedback was given. Contrast of the pattern was varied using a staircase algorithm designed to converge to a 79% correct level (Levitt, 1971). Contrast was decreased after three consecutive correct responses and increased after one wrong response. The size of the contrast increments or decrements decreased as the staircase depth increased,

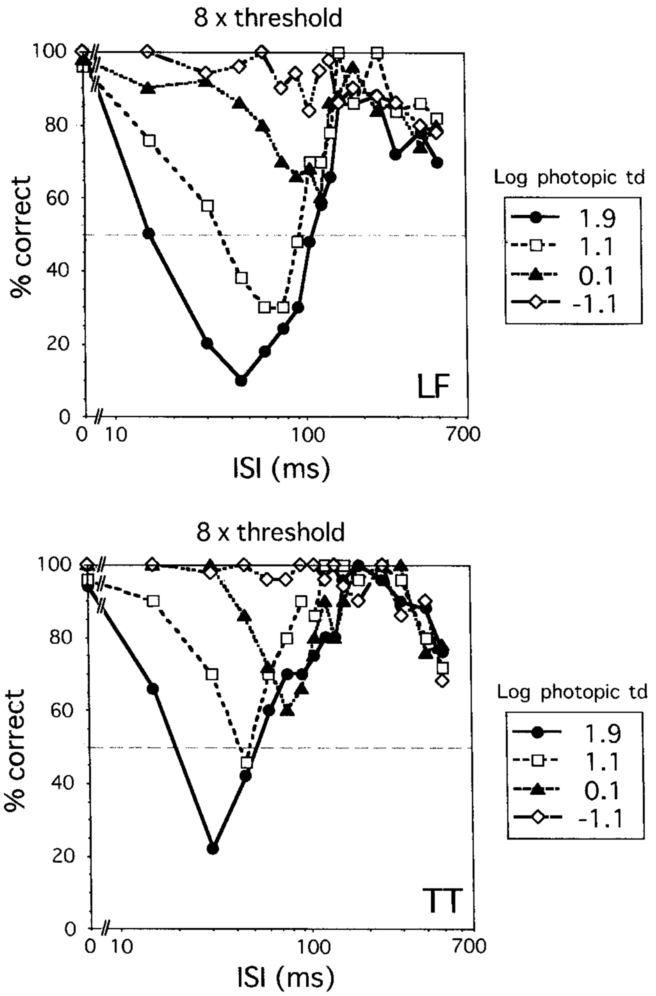

FIGURE 3. The results of Experiment 1 for two subjects (LF, TT). Percent correct response as a function of ISI (msec) is plotted. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction of a two-frame sine-wave grating. When fewer than 50% of the responses were scored as correct, subjects reported apparent motion in the reversed direction on a majority of trials. The stimulus contrast was 8× threshold contrast for direction discrimination. Each curve shows data taken at a different adapting level, from 1.9 log photopic td to −1.1 log photopic td. Note the similarity between this figure and the prediction from the motion energy model shown in Fig. 1(B) when the ISI is short, and the deviation when the ISI is long.

being 0.4 log unit in the beginning and falling to a terminal value of 0.1 log unit. The threshold for a given staircase run was computed as the mean of the contrasts of the final six out of nine turning points. At least three staircases were run to determine each threshold. Similar measurements were made for each subject at each adapting level.

Direction discrimination measurements with ISIs

The method of constant stimuli was used for measurements of direction discrimination. Five hundred milliseconds after the beep signaling the start of each trial, the two-frame sine-wave stimulus described above was displayed. The subject's task was to indicate the direction of motion (upward or downward) by pressing the appropriate button. The button press initiated the next trial. The duration of each of the two frames was 500 msec. A blank field with the same space-averaged luminance as the grating was presented during the ISI. ISIs varied from a nominal 0 to 500 msec. Each session comprised 100 trials presented in random order. Each subject completed at least 40 trials for each ISI value. At each adaptation level, contrasts ranging from 3 to 8 times direction discrimination threshold were used. Contrasts were always referenced to the direction discrimination threshold at the same adaptation level.

Results and discussion

Figure 3 shows results for two subjects and four adapting levels. Spatial frequency was 1.0 c/deg, and contrast was 8× direction discrimination threshold at the corresponding adapting level. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction (short-path direction, in Turano and Pantle's terminology). Thus, when fewer than 50% of the responses are defined as correct, the subject reported apparent motion in the reversed direction on a majority of trials. Both the ISI at which the probability of motion reversal was greatest and the percentage of trials in which the reversal was reported changed systematically as the adapting level changed. Under photopic conditions (1.9 log photopic td), motion reversal was prominent at an ISI of about 45 msec for LF and 30 msec for TT, consistent with the data of Shioiri and Cavanagh (1990) and Pantle and Turano (1992). However, as the adapting level decreased, the ISI at which the strongest motion reversal occurred became longer and the frequency of reversed motion perception decreased.

At the highest adaptation level (1.9 log td), motion reversal was reported on 90% of trials by LF (ISI = 45 msec) and on 78% by TT (ISI = 30 msec). At intermediate adaptation levels, however, the probability of seeing reversed motion decreased markedly at an ISI of 30 or 45 msec, but it changed less at longer ISIs. At the lowest adaptation level (−1.1 log td), at which only the scotopic system is assumed to function, apparent motion reversal completely disappeared, and subjects reported the correct direction at all ISIs. For short ISIs, these results are qualitatively consistent with the motion energy computation shown in Fig. 1(B), which suggested that the strongest motion reversal should occur at longer ISIs, and the absolute frequency of perceiving motion reversal should decrease as the adapting level fell. The disappearance of the illusion at low adapting luminances cannot result either from an inability to detect the stimulus or from an inability to determine the direction of motion, since contrasts were equated in terms of multiples of direction discrimination threshold at each luminance level.

When the ISI was longer than about 100 msec, however, the data deviated from the predictions based upon Fig. 1(B). Under photopic conditions, although the probability of a correct response was significantly below 0.5 at some short ISIs, it gradually increased to approach 1.0 again at an ISI of about 150 msec, falling back toward

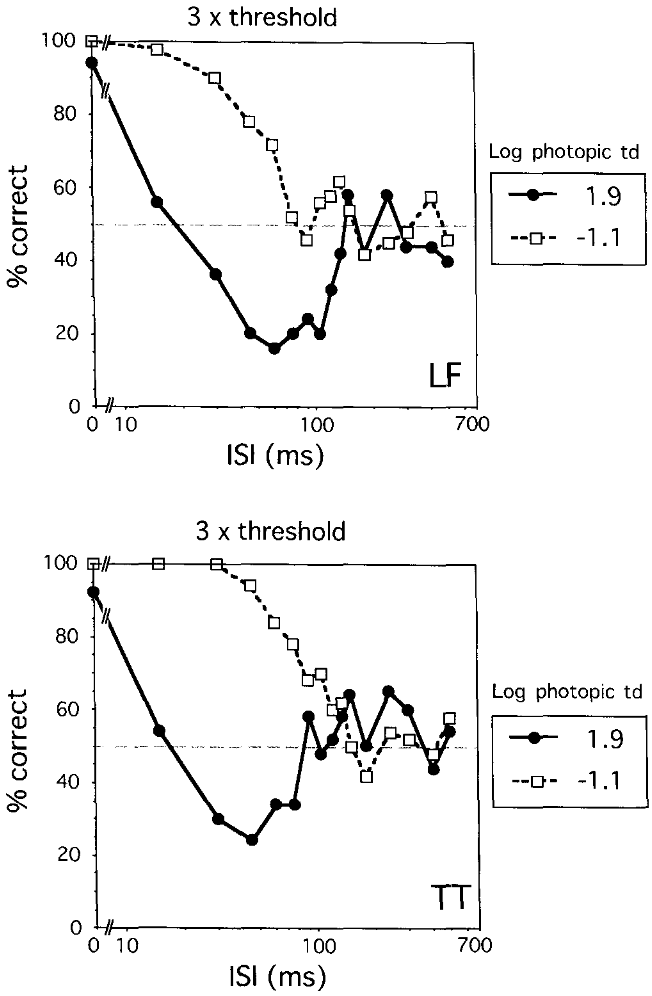

FIGURE 4. The results of Experiment 2 for two subjects (LF, TT). The stimulus contrast was 3x threshold contrast for direction discrimination. Percent correct response as a function of ISI (msec) is plotted. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction of two-frame sine-wave gratings. When fewer than 50% of the responses were scored as correct, subjects reported apparent motion in the reversed direction on a majority of trials. Each curve shows data taken at two adapting levels, 1.9 log photopic and −1.1 log photopic td. Note the similarity between this figure and the prediction from the motion energy model shown in Fig. 1(B) for the full range of ISIs.

chance at ISIs of about 400 msec for LS and 500 msec for TT, though it did not reach chance in the range examined. This behavior was essentially invariant with different adaptation levels. Neither of the first-order motion models predicted this recovery to correct response at intermediate ISIs under photopic conditions. Such a recovery is not seen in the data of Shioiri and Cavanagh (1990), in which there was no tendency to recover the correct response after the reversed motion disappeared. Their subjects' responses remained at chance level for longer ISIs. Recall that their stimuli were random-dot kinematograms. We present a possible explanation for the difference below. Pantle and Turano (1992) did not report data for ISIs longer than 40 msec.

These results suggest that the effect of an ISI of the same mean luminance can be explained by a first-order motion detector, as suggested by Shioiri and Cavanagh (1990) and Pantle and Turano (1992), when the ISI is short. When a longer ISI is inserted, however, our results seem to reflect the operation of another kind of motion system. One candidate is a feature-tracking system that extracts motion information based on the solution of the correspondence problem between prominent features (Ullman, 1979; Dawson, 1991). Although the exact nature of such a system is still unclear, certain motion phenomena can be best explained if we assume that a feature-tracking system is implemented in the visual system (e.g., Georgeson & Harris, 1990; Hammett *et al.*, 1993; Smith, 1994; Lu & Sperling, 1995). The sine-wave grating used in Experiment 1 contained prominent features, bright and dark bars, that could provide a strong input to such a feature tracking system.

Some evidence supports the idea that a feature-tracking system contributes to the deviation from the predictions discussed. First, previous research has suggested that when the ISI is long, only a feature-tracking system is functioning. When a compound grating is stepped (i.e., shown in sampled motion), the perceived direction of motion corresponds to the direction of its strongest Fourier component if the ISI is very short. As the ISI increases, however, the perceived direction shifts to that corresponding to the motion of the grating's spatial features, even when there is no strong corresponding Fourier component (Georgeson & Harris, 1990; Hammett *et al.*, 1993). Secondly, when random-dot kinematograms are used, the effect of varying the ISI is very similar to that predicted from Fig. 1(B) for the photopic condition (transient factor $B = 0.8$) (see Fig. 1 of Shioiri & Cavanagh, 1990). There is no return to correct performance after the disappearance of apparent motion reversal. A feature-matching system presumably does not function well when there are no prominent features that can be easily matched in the motion display, and the perceived direction would then be determined primarily by the output of something like a motion energy detector.

If we assume that a first-order motion system is dominant at short ISIs (we present supporting evidence below) and a feature-matching system is dominant at long ISIs—and there is independent evidence to that effect (Georgeson & Harris, 1990; Hammett *et al.*, 1993; Scott *et al.*, 1993)—then we can explain the results of Experiment 1 at photopic adaptation levels. As the contribution of the first-order motion detector becomes increasingly weaker at longer ISIs, the feature-matching system becomes dominant, producing a recovery to the correct directional percept.

Can the same conjecture be applied to the results under scotopic conditions? Because both the first-order motion system and a simple feature-tracking system would produce qualitatively similar results under these conditions, we cannot conclude from Experiment 1 which system is working at scotopic levels. Experiment 2 was designed to address this question.

EXPERIMENT 2

It has been suggested that the contrast sensitivity of a first-order motion system is relatively higher than that of other motion mechanisms (Sperling, 1989; Dosher *et al.*, 1989; Nishida, 1993; Solomon & Sperling, 1994; Smith *et al.*, 1994; Lu & Sperling, 1995). If two different systems, one first-order mechanism and one feature-tracking mechanism, are actually responding to our motion stimulus, then it might be possible to dissociate them by reducing the stimulus contrast. At a low contrast, only a first-order system might be expected to function if it has higher contrast sensitivity. In Experiment 2, we examined the effect of reducing stimulus contrast under both photopic and scotopic conditions. We suggested above that a first-order system might underlie motion analysis when ISIs are short, and a feature-tracking system might underlie motion analysis when ISIs are long. If so, reducing the contrast of stimuli presented with a long ISI should degrade performance, but it should have little effect on performance when the ISI is short.

Results and discussion

Figure 4 shows the 3× contrast–threshold data from both photopic and scotopic conditions for two subjects. Recall that correct responses are defined as those corresponding to the $\pi/2$ phase shift direction. When fewer than 50% of the responses were "correct" (i.e., in the short-path direction), the subject reported motion reversal on a majority of trials. Under photopic conditions (1.9 log photopic td) at 3× threshold contrast, the maximum likelihood of seeing reversed motion occurred at an ISI of around 50 msec, which is comparable to the results at high contrast (Fig. 3). At the lower contrast, however, there was no recovery to the correct response at intermediate ISIs, although such a recovery was prominent at the higher contrast. Under scotopic conditions (-1.1 log photopic td), the probability of reporting the correct direction of motion of a low-contrast pattern was essentially unity at very short ISIs (ISI $\leq$45 msec). As ISI increased, however, the percentage of correct responses fell to chance more rapidly than at the higher contrast (Fig. 3).

Thus, the *occurrence* of reversed motion appears to be independent of contrast, while the *recovery* to the correct response at long ISIs depends strongly upon stimulus contrast. Note the qualitative similarity between the data from both photopic and scotopic conditions (Fig. 4) and the function shown in Fig. 1(B), in which the output of a first-order motion detector is modeled. This similarity suggests that only a first-order motion mechanism is used at low contrasts under both photopic and scotopic conditions. This further supports our conjecture that something like a higher-order feature-matching mechanism is functioning at the longer ISIs when contrast is high.

Previously, the relative contrast sensitivities of first- and second-order systems have been estimated indirectly, because the definition of contrast differs for first- and second-order stimuli. For example, the carrier depth and the modulation depth of contrast-modulated motion stimuli have been compared with the luminance contrast of first-order stimuli as measured by the usual Michelson relation (Nishida, 1993; Smith *et al.*, 1994). Solomon and Sperling (1994) calculated the relative efficiencies of the mechanisms that detect half-wave, full-wave, and first-order motion, and showed that the efficiencies of non-Fourier systems are relatively lower than that of a first-order system. Lu and Sperling (1995) examined contrast sensitivity for the motion of a luminance-modulated pattern and found sensitivity to be higher under monocular viewing conditions than for interocular viewing. Based upon their conclusion (from other experiments) that first- and second-order systems are exclusively monocular, they argued that only a feature-tracking system could be operating under interocular viewing, and that the contrast sensitivity of such a system is lower than that of first- and second-order systems. Our results show directly that a feature-tracking system has lower contrast sensitivity than a first-order motion system.

EXPERIMENT 3

Experiment 3 was designed to test the conjecture that the shape of the temporal impulse response function is critical for the occurrence of reversed motion. Previous research has shown that the temporal impulse response function for isoluminant stimuli is essentially monophasic, with little or no negative lobe (Smith *et al.*, 1984; Uchikawa & Ikeda, 1986; Swanson *et al.*, 1987; Burr & Morrone, 1993). Burr and Morrone (1993) used a two pulse resolution method and showed that the derived temporal impulse function can explain temporal contrast sensitivity for the detection of counterphase-modulated gratings. They also suggested that the monophasic character of the temporal impulse response function for isoluminant stimuli would be important in understanding motion detection at isoluminance. If the monophasic characteristic of the temporal impulse response function is the cause of the reversed motion perception, then reversed motion should not appear when a color-defined grating is used as the moving stimulus.

A monophasic temporal impulse response function is also found for high spatial frequency luminance gratings under photopic conditions. Watson and Nachmias (1977) measured two-pulse resolution for luminance gratings and estimated the temporal impulse response function for several spatial frequencies. They found that the negative lobe of the temporal impulse response function is very small at 7.0 c/deg, and absent at 10.5 c/deg. As with isoluminant stimuli, then, reversed motion should disappear when the stimulus is a luminance-defined grating of high spatial frequency.

In Experiment 3, we tested whether a color-defined grating or a luminance-defined high spatial frequency grating would induce the perception of reversed motion.

Procedure

In order to minimize longitudinal chromatic aberra-

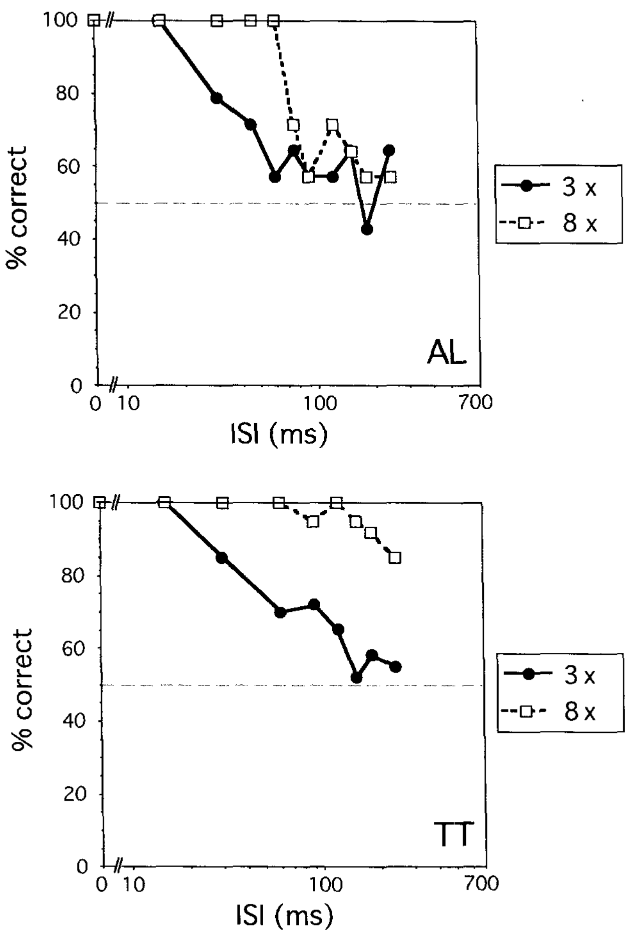

FIGURE 5. Results for two subjects (AL, TT) from Experiment 3, in which stimuli were isoluminant gratings of 1 c/deg. Percent correct response is plotted as a function of ISI (msec) of the two-frame sine-wave gratings. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction. When fewer than 50% of the responses were scored as correct, subjects reported apparent motion in the reversed direction on a majority of trials. Two contrasts (3× and 8× contrast threshold for direction discrimination) were used.

tion, subjects viewed the stimulus monocularly through an air-spaced achromatizing lens (Powell, 1981). Precise alignment was achieved by having the subject adjust his/her position using a two-dimensional red and blue vernier target (Kooi & De Valois, 1992). To maintain correct alignment, a bite bar mounted on an XYZ translator was used. Color contrast was produced by modulation along a red–cyan color axis (CIE x, y for red = 0.603, 0.352, cyan = 0.189, 0.300 and white = 0.307, 0.330). The adaptation level was 1.9 log photopic td. Equal sensation luminances were determined for each subject using heterochromatic flicker photometry at 16.5 Hz. Contrast sensitivity and direction discrimination measurements were as described in Experiment 1. The spatial frequency of the chromatic grating was 1.0 c/deg.

For luminance-defined gratings of high spatial frequency (8.0 c/deg), procedures were identical to those in Experiment 1. A 2 mm artificial pupil was used, and the adapting level was set to 1.9 log photopic td. To reduce any possible effects of the spatial quantization error associated with a video display, viewing distance was set to 230 cm. One degree of visual angle contained 160 pixels at this viewing distance.

Results and discussion

Figure 5 shows results for the color grating for two subjects. Contrasts were 3× or 8× direction discrimination threshold. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction. When fewer than 50% of responses were coded as correct, motion reversal was reported on a majority of trials. With isoluminant stimuli, reversed motion did not appear at either contrast. This is consistent with the idea that the negative lobe of the temporal impulse response function is responsible for the reversed motion perception. As ISI increased, the percentage of correct responses rapidly fell to chance at the low contrast, which is consistent with the prediction from the output of the motion energy model shown in Fig. 1(B). When the contrast was high, correct responses were produced at the longer ISIs, suggesting the contribution of a feature-tracking system. This contrast dependency of the effect of various ISIs is

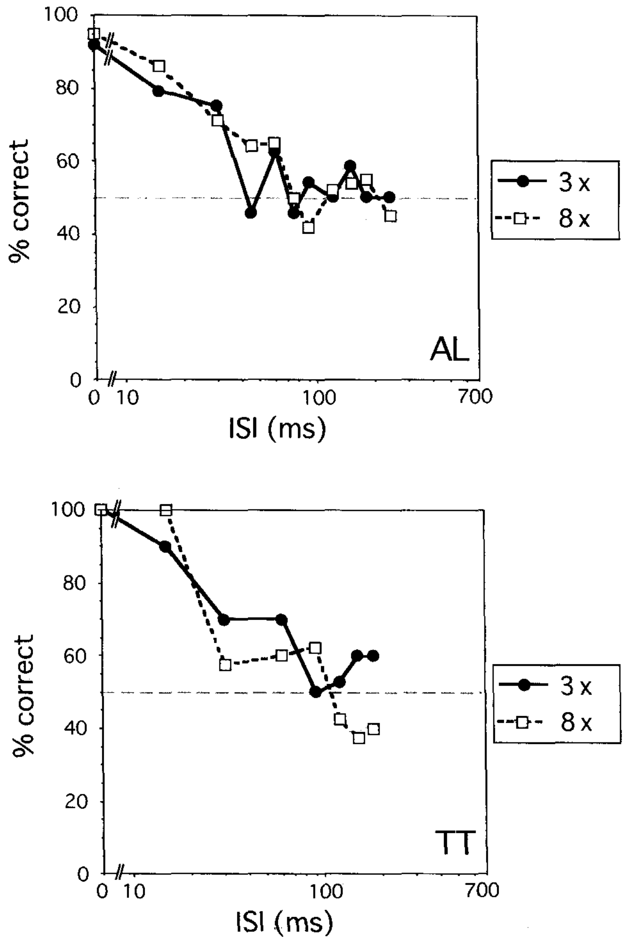

FIGURE 6. The results of Experiment 3 using luminance-varying gratings of 8.0 c/deg for two subjects (AL, TT). Percent correct response is plotted as a function of ISI (msec) of two-frame sine-wave gratings. Correct responses are defined as those corresponding to the $\pi/2$ phase shift direction. When fewer than 50% of the responses were scored as correct, subjects reported apparent motion in the reversed direction on a majority of trials. Two contrasts (3× and 8× contrast threshold for direction discrimination) were used.

similar to that observed for low spatial frequency luminance gratings under scotopic conditions (−1.1 log photopic td data in Figs 3 and 4). Although the role of color in the motion system is still controversial and beyond the scope of this paper, this result seems to suggest that both a motion energy system and a feature-tracking mechanism can function in coding the direction of motion of isoluminant color-defined gratings.

Results for the 8.0 c/deg luminance-defined grating are shown in Fig. 6. Correct responses, again defined as those corresponding to the $\pi/2$ phase shift direction, are plotted as a function of ISI. Pantle and Turano (1992) showed that reversed motion occurred when a 4.0 c/deg luminance-defined grating was used, though the strongest illusion appeared at longer ISIs than found with a 1.0 c/deg grating. In this experiment, we found that a luminance grating of 8 c/deg failed to elicit reversed motion, and the probability of correct direction discrimination decreased monotonically as ISI increased. This result further supports the idea that the monophasic temporal impulse response function contributes to the production of reversed motion. No clear contrast dependence was found for either subject when high spatial frequency gratings were used (Fig. 6). This result can be understood if we assume that a feature-tracking system fails for the high spatial frequency stimulus used in this experiment.

As previously noted, Shioiri and Cavanagh's (1990) results are qualitatively consistent with the output of a motion energy model. A one-dimensional high spatial frequency grating and a field of random dots are similar in that neither contains prominent local features that can be easily tracked. Thus, prominent features, which are present in the low spatial frequency grating stimuli but not in either the random-dot stimuli of Shioiri and Cavanagh (1990) or the high frequency grating we used, may be the key to producing correct responses at longer ISIs (Fig. 6).

GENERAL DISCUSSION

We used a simple display in which two frames of a sine-wave grating were separated by various ISIs equated in mean luminance level. In Experiment 1 we found that the perceived direction of displacement varied with the values of ISI and the adapting level. As adapting level decreased, apparent motion reversal disappeared. This result can be predicted from the output of a first-order motion detector if we assume that the shape of the underlying temporal impulse response function varies with adapting level, from biphasic under photopic conditions to monophasic under scotopic conditions. However, when the ISI was long, results deviated from the prediction derived from a first-order motion detector. In Experiment 2, we showed that this deviation depends upon the stimulus contrast, suggesting that a feature-tracking mechanism, whose contrast sensitivity is assumed to be lower than that of a first-order motion mechanism, is responsible for the failure of prediction. These two experiments suggest that at least two different kinds of motion mechanisms, a first-order motion mechanism and a feature-tracking mechanism, function under both photopic and scotopic adaptation levels. In Experiment 3, we presented additional evidence that the shape of the temporal impulse response function affects the perceived direction of motion. For both isoluminant and high spatial frequency gratings, for which the temporal impulse response has no inhibitory region, no motion reversal was observed.

Our finding of correct (i.e., not reversed) direction perception at all ISIs under scotopic conditions is explained well by assuming that the underlying temporal impulse response function is monophasic if a first-order motion detector is working at low adaptation levels. This conjecture can be checked by different psychophysical tasks than those that we used. For example, it is known that the motion aftereffect (MAE) with a static test pattern occurs only when a subject adapts to a first-order motion stimulus. Second-order stimuli and long-range apparent motion stimuli fail to produce a strong MAE when static test patterns are used (Anstis, 1980; Nishida & Sato, 1995). In informal observations we confirmed that an MAE is produced when static test stimuli are used under scotopic conditions. This is consistent with our conjecture that a first-order motion detector is operative under scotopic conditions.

In this research we were particularly concerned with temporal response properties at various adaptation levels, and the resulting motion perception. Although the visual spatial domain under scotopic conditions is beyond the scope of this study, we should mention the characteristics of the spatial filters at low adaptation levels. The spatial impulse response function used in current first-order motion models is an analog of the receptive fields of simple cells in V1 (Adelson and Bergen, 1985; Watson & Ahumada, 1985; Nowlan & Sejnowski, 1994). There is some evidence that receptive field sizes change as adaptation level changes (e.g., Ramoa *et al.*, 1985). The psychophysically measured contrast sensitivity function (CSF) for scotopic vision differs from the photopic CSF in showing a high spatial frequency cut-off at lower spatial frequencies, a loss of low spatial frequency attenuation, and generally lower contrast sensitivity for all spatial frequencies (e.g., De Valois *et al.*, 1974). This high spatial frequency attenuation results in a derived spatial impulse response function with a smaller inhibitory lobe. The nature of the spatial filters at the front end of a first-order motion detector is still controversial (e.g.; Yang & Blake, 1994), and further research is needed to clarify their structure.

Current psychophysical research suggests that several types of motion detectors are implemented in the visual system (e.g., Smith, 1994; Lu & Sperling, 1995). Our results argue that at least two different motion mechanisms, a first-order motion detector and a feature-tracking mechanism, underlie the direction discrimination of a two-frame sine-wave grating. Some motion phenomena, however, can be well explained by a non-linear operation such as rectification, which is followed by a motion mechanism like that proposed for first-order motion detectors. For these phenomena, it is not necessary to postulate feature-tracking mechanisms (e.g., Chubb & Sperling, 1988; Wilson *et al.*, 1992). Pantle and Turano (1992) showed that apparent motion reversal did not occur when a second-order, contrast-modulated pattern was used, although they only examined ISIs of 50 msec or less. This suggests that the temporal impulse response function of a second-order motion detector that can extract motion information from a contrast modulated pattern may be monophasic, if indeed it does feed into a first-order-like mechanism following its non-linear stage.

In Experiment 3, we suggested that a first-order motion system underlies the analysis of color-defined motion at low contrast, and a feature-tracking system becomes relatively more effective as contrast increases. Cropper and Derrington (1994) measured the minimum velocity required to discriminate the direction of motion, and found that the direction of high contrast color-defined gratings was discriminable at shorter (by factor of ten) presentation durations than that of low contrast chromatic gratings. From these results, they have suggested that the motion of color-defined gratings is detected by a first-order motion system when contrast is high, and by a second-order motion system (a slow-acting system) when contrast is low (see also Derrington & Henning, 1993). De Valois and Bullimore (1992) measured the minimum displacement threshold for a two-frame Gabor patch and found different contrast response functions for luminance variations and color variations. For luminance defined patches, displacement thresholds were invariant for contrast levels greater than about 4× detection threshold. For the color-defined patches, however, displacement thresholds continued to decrease as contrast increased. They suggested that different kinds of motion detection systems function, one a first-order motion mechanism for luminance stimuli, and the other a positional mechanism, in which the position of features is explicitly coded for color stimuli. The reason for the differences between these experiments and the present study is unclear, and further research is needed to clarify how different motion systems function under isoluminant conditions.

The directionally selective simple cells of V1 generally work as a quasi-linear system, and have space–time oriented receptive fields (McLean & Palmer, 1989; DeAngelis *et al.*, 1993). Emerson *et al.* (1992) suggested that Adelson and Bergen's motion energy model is implemented in complex cells in V1. From a computational point of view, the operation required to accomplish feature-tracking is quite different from that of the extraction of orientation in the space–time domain. In a feature-tracking mechanism, prominent features are spatially localized, and corresponding features must be matched across time. Direct physiological evidence of the computation of feature matching has not been reported, although Dawson (1991) suggested the posterior parietal cortex as a candidate neural site. A feature-tracking mechanism would require several distinct stages to solve the correspondence problem. If the quasi-linear first-order motion detector is instantiated in V1, it is reasonable to think that a feature-tracking mechanism probably occurs later in the system.

REFERENCES

Adelson, E. H. & Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2*, 284–299.

Anstis, S. M. (1980). The perception of apparent motion. *Philosophical Transactions of the Royal Society of London B, 290*, 153–168.

Boulton, J. & Baker, C., Jr. (1993). Dependence on stimulus onset asynchrony in apparent motion. Evidence for two mechanisms. *Vision Research, 33*, 2013–2019.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B, 290*, 137–151.

Burr, D. C. & Morrone, M. C. (1993). Impulse–response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America A, 10*, 1706–1713.

Burr, D. C., Ross, J. & Morrone, M. C. (1986). Seeing objects in motion. *Proceedings of the Royal Society of London B, 227*, 249–265.

Cavanagh, P. & Mather, G. (1989). Motion: the long and short of it. *Spatial Vision, 4*, 103–129.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986–2007.

Cropper, S. J. & Derrington, A. M. (1994). Motion of chromatic stimuli: first-order or second-order? *Vision Research, 34*, 49–58.

Dawson, M. R. (1991). The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychological Review, 98*, 569–603.

DeAngelis, G. C., Ohzawa, I. & Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology, 69*, 1118–1134.

Derrington, A. M. & Henning, G. B. (1993). Detecting and discriminating the direction of motion of luminance and colour gratings. *Vision Research, 33*, 799–811.

De Valois, K. K. & Bullimore, M. A. (1992). Displacement detection for luminance and isoluminant Gabor patches. *Investigative Ophthalmology and Visual Science, 33*, 972.

De Valois, R. L. & De Valois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

De Valois, R. L., Morgan, H. & Snodderly, D. M. (1974). Psychophysical studies of monkey vision—III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research, 14*, 75–81.

Dosher, B. A., Landy, M. S. & Sperling, G. (1989). Kinetic depth effect and optic flow—I. 3D shape from Fourier motion. *Vision Research, 29*, 1789–1813.

Emerson, R. C., Bergen, J. R. & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research, 32*, 203–218.

Georgeson, M. A. & Harris, M. G. (1990). The temporal range of motion sensing and motion perception. *Vision Research, 30*, 615–619.

Hammett, S. T., Ledgeway, T. & Smith, A. T. (1993). Transparent motion from feature- and luminance-based processes. *Vision Research, 33*, 1119–1122.

Hecht, S. & Shlaer, S. (1936). Intermittent stimulation by light—V. The relation between intensity and critical frequency for different parts of the spectrum. *Journal of General Physiology, 19*, 965–979.

Hood, D. C. & Finkelstein, M. A. (1986). Visual sensitivity. In Boff, K., Kaufman, L. & Thomas, J. (Eds), *Handbook of perception and human performance* (Vol. 1, Chap. 5, pp. 1–66). New York: Wiley.

Ikeda, M. (1986). Temporal impulse response. *Vision Research, 26*, 1431–1440.

Kelly, D. H. (1971). Theory of flicker and transient responses—I.

Uniform fields. *Journal of the Optical Society of America, 61*, 537–546.
Kooi, F. L. & De Valois, K. K. (1992). The role of color in the motion system. *Vision Research, 32*, 657–668.
Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.
Lu, Z.-L. & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*, 2697–2722.
Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B, 211*, 151–180.
McCourt, M. E. (1990). Disappearance of grating induction at scotopic luminances. *Vision Research, 30*, 431–437.
McLean, J. & Palmer, L. A. (1989). Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in the cat's striate cortex. *Vision Research, 29*, 675–679.
Nakayama, K. & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America A, 2*, 267–274.
Nishida, S. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research, 33*, 633–645.
Nishida, S. & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Research, 35*, 477–490.
Nowlan, S. J. & Sejnowski, T. J. (1994). Filter selection model for motion segmentation and velocity integration. *Journal of the Optical Society of America A, 11*, 3177–3200.
Pantle, A. & Turano, K. (1992). Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. *Vision Research, 32*, 2093–2106.
Powell, I. (1981). Lenses for correcting chromatic aberration of the eye. *Applied Optics, 20*, 4152–4155.
Ramachandran, V. S. & Cavanagh, P. (1987). Motion capture anisotropy. *Vision Research, 27*, 97–106.
Ramoa, A. S., Freeman, R. D. & Macy, A. (1985). Comparison of response properties of cells in the cat's visual cortex at high and low luminance level. *Journal of Neurophysiology, 54*, 61–72.
Robson, J. G. (1966). Spatio and temporal contrast-sensitivity function of the visual system. *Journal of the Optical Society of America, 56*, 1141–1142.
Scott, J. A., De Valois, K. K. & Bullimore, M. A. (1993). The influence of interstimulus interval on displacement discrimination. *Investigative Ophthalmology and Visual Science, 34*, 1035.
Shioiri, S. & Cavanagh, P. (1990). ISI produces reverse apparent motion. *Vision Research, 30*, 757–768.
Snowden, R. J., Hess, R. F. & Waugh, S. J. (1995). The processing of temporal modulation at different levels of retinal illuminance. *Vision Research, 35*, 775–789.
Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America A, 11*, 1940–1948.
Smith, A. T., Hess, R. F. & Baker, C.Jr. (1994). Direction identification thresholds for second-order motion in central and peripheral vision. *Journal of the Optical Society of America A, 11*, 506–514.
Smith, V. C., Bowen, R. W. & Pokorny, J. (1984). Threshold temporal integration of chromatic stimuli. *Vision Research, 24*, 653–660.
Solomon, J. A. & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research, 34*, 2239–2257.
Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision, 4*, 183–207.
Strout, J. J., Pantle, A. & Mills, S. L. (1994). An energy model of interframe interval effects in single-step apparent motion. *Vision Research, 34*, 3223–3240.
Swanson, W. H., Ueno, T., Smith, V. C. & Pokorny, J. (1987). Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. *Journal of the Optical Society of America A, 4*, 1992–2005.
Turano, K. & Pantle, A. (1985). Discontinuity limits for the generation of visual motion aftereffects with sine- and square-wave gratings. *Journal of the Optical Society of America A, 2*, 260–266.
Uchikawa, K. & Ikeda, M. (1986). Temporal integration of chromatic double pulses for detection of equal-luminance wavelength changes. *Journal of the Optical Society of America A, 3*, 2109–2115.
Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
van Santen, J. P. H. & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A, 1*, 451–473.
van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2*, 300–321.
Watson, A. B. (1986). Temporal sensitivity. In Boff, K., Kaufman, L. & Thomas, J. (Eds), *Handbook of perception and human performance* (Vol. 1, Chap. 6, pp. 1–42). New York: Wiley.
Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2*, 322–341.
Watson, A. B. & Nachmias, J. (1977). Patterns of temporal interaction in the detection of gratings. *Vision Research, 17*, 893–902.
Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.
Yang, Y. & Blake, R. (1994). Broad tuning for spatial frequency of neural mechanisms underlying visual perception of coherent motion. *Nature, 371*, 793–796.

Acknowledgements—This research was supported by grants from the National Science Foundation (IBN 9409310) and the National Eye Institute (EY00014). T. Takeuchi was supported by the Japan Society for the Promotion of Science.